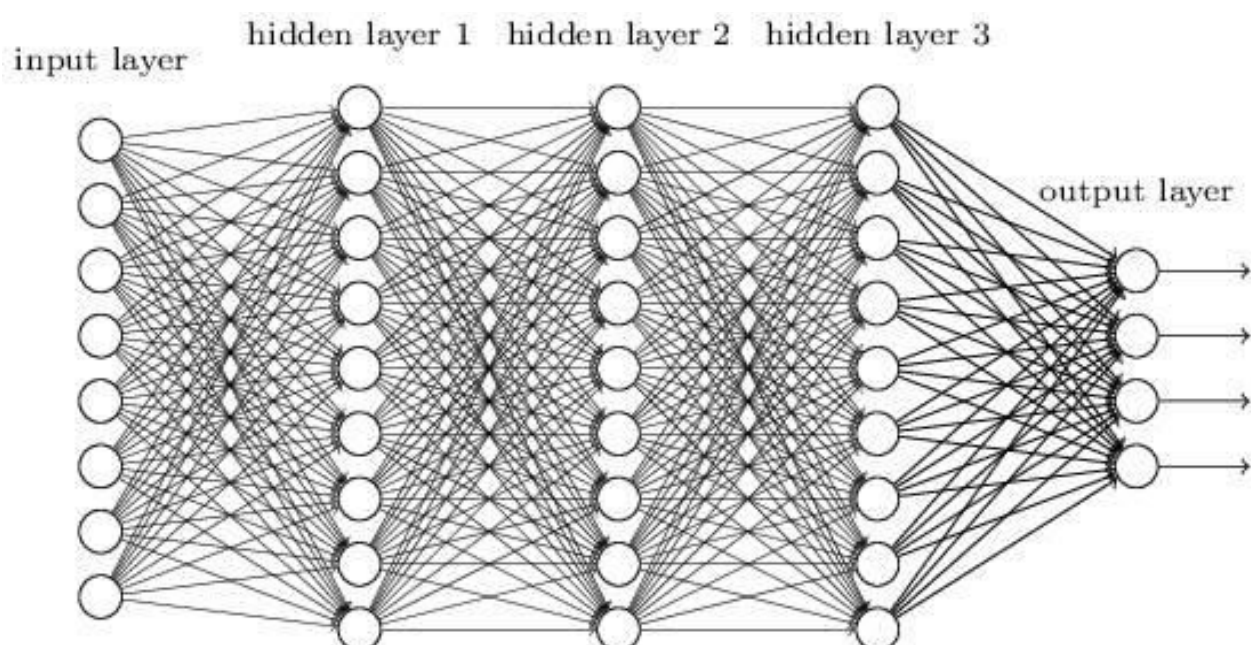

网络改进 & 参数调整

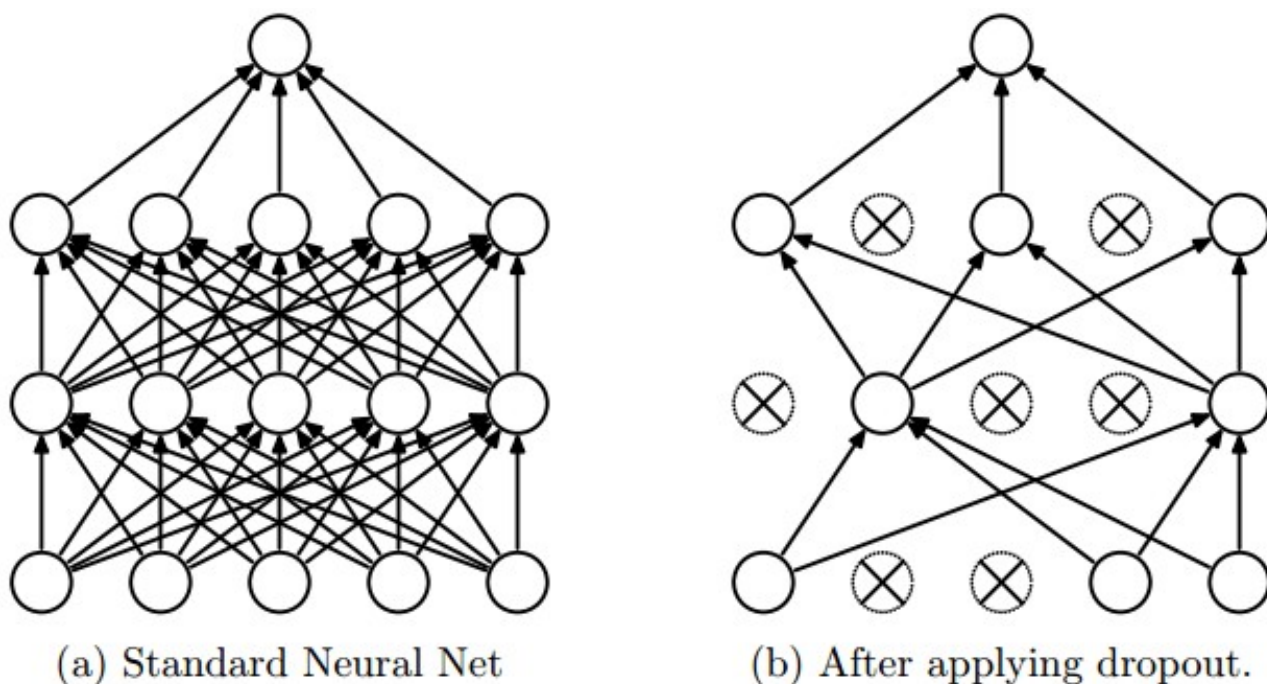
耿同欣 - 2017年10月28日



dropout

原理

dropout是一种防止模型过拟合的技术，它的基本思想是在训练的时候随机的dropout(丢弃)一些神经元的激活，这样可以让模型更鲁棒，因为它不会太依赖某些局部的特征（因为局部特征有可能被丢弃）。



上图 (a) 是标准的一个全连接的神经网络，(b) 是对 (a) 应用了 dropout 的结果，它会以一定的概率（ dropout probability）随机的丢弃掉一些神经元。

引自<http://geek.csdn.net/news/detail/161276>

实现

开始，我将 keep_prob 参数设置为 0.8，即样本有 80% 被保留下来，但经过测试仍会出现过拟合现象

```
# 训练
r = 201 * 14 / batch_size + 1
for i in range(200):
    for j in range(r):
        batch = get_batch(j)
        train_step.run(feed_dict={x: batch[0], y_: batch[1], keep_prob: 0.8})
    print "%d - 正确率: %g" % (i, accuracy.eval(feed_dict={x: getTest()[0], y_: getTest()[1], keep_prob: 1.0}))
    if i > 80 and (float(accuracy.eval(feed_dict={x: getTest()[0], y_: getTest()[1], keep_prob: 1.0})) > 0.988):
        break
```

适当调低keep_prob 参数，舍弃更多样本，继续测试。

将 keep_prob 参数设置为 0.5，即样本有 50% 被保留下来，有 50% 被舍弃，迭代 200 次时，正确率稳定在 98.3 % - 98.5 %。

学习率调整

开始我将 learning rate 设置为 $2 * 1e-3$ ，正确率在升到最高后开始急剧下降。经查阅资料这是过拟合的表现，可通过减小 learning rate 的方法解决。

试着调小 learning rate 到 $1e-3$ ，并且设置当连续4次以上正确率在 98.7%以上时（或迭代 300 次后）停止训练。

最好的模型验证集准确率为98.83%。